

ARTICLE

What's in, what's out? Towards a rigorous definition of the boundaries of benefit-cost analysis

Daniel Acland 

Goldman School of Public Policy, University of California Berkeley, 2607 Hearst Ave, Berkeley, CA 94720, USA
Email: acland@berkeley.edu

(Received 10 January 2019; revised 13 July 2020; accepted 27 October 2020; first published online 24 February 2021)

Abstract

Benefit-cost analysis (BCA) is typically defined as an implementation of the potential Pareto criterion, which requires inclusion of any impact for which individuals have willingness to pay (WTP). This definition is incompatible with the exclusion of impacts such as rights and distributional concerns, for which individuals do have WTP. I propose a new definition: BCA should include only impacts for which consumer sovereignty should govern. This is because WTP implicitly preserves consumer sovereignty, and is thus only appropriate for 'sovereignty-warranting' impacts. I compare the high cost of including non-sovereignty-warranting impacts to the relatively low cost of excluding sovereignty-warranting impacts.

Keywords: benefit-cost analysis; consumer sovereignty; willingness to pay; welfare

1. Introduction

The standard definition of benefit-cost analysis (BCA) is that it is an implementation of the potential Pareto criterion (PPC) – also known as the Kaldor-Hicks criterion – for determining the relative efficiency of different policies. The PPC states that a policy should be implemented if those who gain from the policy could compensate those who lose and still be made better off. The net benefit of a policy, then, relative to the status quo, is the difference between the maximum that the 'winners' of the policy would be willing to pay for the positive impacts they experience, and the minimum that the 'losers' would be willing to accept in exchange for the negative impacts they experience. In other words, it is the surplus the winners would be left with if they compensated the losers.¹

¹I realize that there are numerous caveats, including the discrepancy between willingness to pay and willingness to accept, the problem of socially unacceptable preferences, the numerous problems of irrationality and inattention that have been raised by behavioural economists, and others. Without, I

To call this the textbook definition of BCA is literally correct. Boardman *et al.* (2011), arguably the standard BCA textbook, states, ‘CBA utilizes . . . the *Kaldor-Hicks criterion* . . . the basis for the *potential Pareto efficiency rule* . . .’ (emphasis theirs). Edward Gramlich’s (1990) textbook states that the fundamental principle of BCA is ‘to choose [the program] that maximizes net social benefits,’ and elsewhere clarifies, ‘net national welfare [presumably synonymous with net social benefit] can be maximized by following . . . the Kaldor-Hicks rule’.

Implicit in this definition is the requirement that any policy impact individuals would be willing to pay to gain, or be paid to accept losing, must be included in BCA, if methodologically possible. Because if not, we haven’t tallied the total amount the winners would be willing to pay, or the losers would need to be paid, and thus don’t have a full accounting of net benefit. If BCA is defined in terms of the PPC, the requirement that any impact be included for which there is willingness to pay is not subject to debate.

And yet, there are policy impacts that are widely thought to not belong in BCA. Boardman *et al.* cite equality of opportunity and equality of outcome. *Circular A-4*, the primary guidance document for regulatory impact assessment within the federal government of the United States, includes distributional effects, discrimination, privacy, personal freedom, and ‘other democratic aspirations’ (Office of Management and Budget 2003). Adler and Posner (2006) include, among the reasons an individual might prefer one policy to another, but which should not be included in BCA, ‘moral reasons, detached sympathy . . . national or ethnic reasons . . . and so on’.

If it could be shown that there are individuals who have willingness to pay (WTP) for the impacts that are typically thought to not belong in BCA, we would have evidence of a conflict between two ways of thinking about what should be included in BCA. And, importantly, this would not be a conflict between opposing groups. It would be a conflict between two ideas that the broad consensus within the field of BCA hold to be true. It would mean that the large majority of scholars and practitioners within the field believe both that BCA is defined by the PPC, which implies that everything for which there is WTP must be included, *and* that BCA is *not* defined by the PPC, because it should not include certain things for which there is WTP. This would be a fundamental contradiction.

I have used contingent valuation surveys to elicit WTP for impacts of the kind that we typically think should not be included in BCA. Here is the list of the impacts for which I elicited WTP, and, in parentheses, the values I believe are captured in individuals’ WTP for them: (1) preservation of an indigenous culture (justice, fairness, dignity), (2) subsidization of long-acting reversible contraceptives (right to be free of unwanted outcomes, general sense of a good society), (3) reduction in income inequality (altruism, fairness, social cohesion), (4) allowing or banning transgender individuals from the military (rights, justice, altruism, out-group animus). I infer the values in parentheses from questions in my surveys about the reasons individuals had for stating a willingness to pay for each of the impacts. Thus, I not only have empirical evidence of WTP, but also empirical

believe, losing the validity of my argument, I am going to ignore all of these caveats, purely for the purpose of simplicity.

evidence of what my respondents were evincing WTP for. These results demonstrate that there *are* policy impacts that satisfy the PPC definition of what should be included in BCA (individuals have WTP), while also satisfying consensus opinion about what should *not* be included. This brings to a head the question of what should and should not be included in BCA, which is to say, the definition of BCA.

Zerbe (2004) has proposed one solution, which is to simply double down on the PPC: 'My suggestion . . . is that all goods should be included in economic analysis for which there is a willingness to pay (WTP).' This view has been reiterated by David Weimer, one of the coauthors of the Boardman text book, in a private communication (2017): 'I see a good as anything someone is willing to pay to obtain (or avoid).' If this is to be the accepted definition, then any list of impacts that we feel ought *not* to be included should be subjected to the test of whether individuals have WTP for them, and if individuals do have WTP for them, there should be no further mention of not including them. On the basis of my contingent valuation surveys, most of the impacts excluded by consensus would have to be included. This would be a dramatic change in the consensus approach to BCA.

An alternative approach would be to set aside the PPC and adopt a new definition of BCA. The first step would be to clarify what BCA is. To begin, BCA either is, or is not a decision rule. A decision rule is a single test that determines whether a policy should be adopted. As a decision rule, BCA states that a policy should be adopted if it increases net benefit relative to the status quo, or maximizes net benefit relative to a set of other alternatives, and not otherwise. In other words, the decision rule is the PPC. Now, if BCA is a decision rule, based on the PPC, then clearly every impact for which there is WTP must be included, because otherwise the rule has not been comprehensively applied. But, despite the fact that both the Executive Order that first established BCA as a requirement for federal regulations in the USA, and all of the guidance documents on BCA published by the federal government, implicitly insist that BCA should be used as a decision rule, there is no justification for this.

Three attempts have been made to justify BCA as a decision rule. The first sees treating BCA as a decision rule as a way to (at least approximately) base policy decisions on strict utilitarianism, in which decisions are made on the basis of whether they increase, on net, something like wellbeing, or happiness, or satisfaction.² The problem with this justification is that utilitarianism has been decisively rejected as a moral foundation for public decision making, for a host of reasons (Kelman 1981). Other values simply must be taken into account. Two other justifications are based on the idea of BCA as an implementation of the Kaldor-Hicks criterion, both arguing that real-world application of BCA as a decision rule can be thought of as an approximate implementation of the less controversial Pareto criterion – some are made better off while no one is made worse off. One justification claims that this *will* be so, because over a large number of policies decided by BCA, the losers in some cases will be the winners

²To be properly applied, this approach requires that WTP be weighted by the marginal utility of wealth to overcome bias against the poor, an issue I address in section 3.

in others (Hotelling 1938; Hicks 1941). The other claims that using BCA as a decision rule *could* achieve Pareto efficiency, because it ensures that the winners could pay off the losers, and thus allows for true Pareto improvements using ex-post redistribution (Kaldor 1939), leaving the ultimate implementation of the Pareto criterion in the hands of politicians. Both of these justifications are problematic because neither the balancing out over time, nor the redistribution after the fact, actually happens, leaving BCA, as a decision rule, merely an implementation of the Kaldor-Hicks criterion, which has been criticized on many grounds, both practical and philosophical, and has largely been rejected as a moral foundation for decision making.³

It seems, then, that BCA cannot be seen as a decision rule. Instead, it should be seen as a measure of one kind of impact of a policy, to be weighed by decision makers against other impacts.⁴ Adler and Posner (2006) take this position, saying that what BCA measures is relevant to decision making, but must be set alongside 'a plurality of moral factors'. The question that needs to be addressed is, if BCA is not a decision rule, but only a source of information about one category of impacts, what is that category? In other words, what is the definition of BCA? In my view, the answer begins with an understanding of the kind of information BCA contains, which is to say, the kind of information contained in WTP. Once the nature of the information contained in BCA is understood, the approach to defining what should be included in BCA will become clear.

2. What kind of information does BCA contain?

Adler and Posner, in *New Foundations for Cost-Benefit Analysis* (2006), have proposed one way to differentiate between included and excluded impacts, and in the process have determined the kind of information they feel BCA should contain, and the type of impacts for which that kind of information is appropriate. They begin with the conventional understanding that BCA should be a measure of welfare, and should not include impacts that do not contribute to welfare. This requires them to define welfare, which they do by starting with what is called a 'preferentialist' account of welfare, and then defining two types of preferences, 'disinterested' and 'self-interested'. In defining disinterested preferences they write, 'people can prefer states that are, intuitively, unrelated to their own lives'. And they include among those states the list I referenced earlier. Self-interested preferences are those that are for states that are directly

³For a review of criticisms, see Bostani and Malekpoor (2012).

⁴It is perhaps not surprising that the federal government of the United States, in its executive orders and guidance documents, is unclear as to whether or not BCA is to be used as a decision rule. Executive Order 12291, 3 CFR: 127. 1981, which established the requirement that all major federal regulations be subjected to regulatory impact assessment, states that, 'Regulatory action shall not be undertaken unless the potential benefits to society for the regulation outweigh the potential costs to society', and, 'Regulatory objectives shall be chosen to maximize the net benefits to society'. This unambiguously implies that BCA is a decision rule. Meanwhile, Circular A-4, the primary guidance document on BCA for the federal government, states, 'benefit-cost analysis provides decision makers with a clear indication of the most efficient alternative, that is, the alternative that generates the largest net benefits to society (ignoring distributional effects)'. This implies that BCA is not to be used as a decision rule. Distributional effects can also be considered.

related to individuals' lives. They then define welfare as the sum of all of an individual's self-interested preferences. And thus they argue that BCA should include only the value individuals place on their self-interested preferences for the states (impacts) generated by a policy. With this welfare-based definition they argue that they have placed BCA on a 'new foundation', ensuring that it remains a measure of individual welfare, or wellbeing, rather than simply a measure of the aggregate dollar amount individuals would be willing to sacrifice for the satisfaction of all of their preferences.⁵ And implicitly, they believe that individual welfare is the value for which BCA conveys the appropriate kind of information.

My approach to inclusion and exclusion is importantly different. I begin by asking, 'What is the nature of the information conveyed in BCA or in other words, in WTP?' For any given impact, it is the value individuals place on getting or not getting that impact, measured in the dollar value of other goods that would generate the same satisfaction as that, whether it be the gain or loss of a piece of consumer electronics, or of a certain quantity of airborne particulates, or of a feeling of living in a civilized society. Because we all engage in monetary trades all the time, we intuitively understand what it means to say that a smartphone generates \$1,000 of value for an individual, and that if the same individual would be willing to pay \$1,000 for a given reduction in airborne particulates, the phone and the pollution reduction are of equal value to the individual.⁶ What is left *out* of WTP is any information about *why* individuals value the impact, or the *kind* of value they receive from it. And WTP contains no information about why the impact might matter to society as a whole, over and above the value it generates for the individuals receiving it. Thus, in my view, WTP provides appropriate information only for impacts for which we feel (i) that it should not matter to society *why* individuals value them, and (ii) for which we feel there are no other reasons they matter to society, over and above the fact that they generate value for the individuals who receive them. This is the basis on which we should decide what to include in BCA.

For example, if a policy has the effect of allowing an individual to purchase a smartphone (by removing a tariff, say), the value to society of that smartphone is purely the value it generates for the individual who gets to use it (plus or minus the opportunity cost of the resources used in its production, and any externalities its production or use generate to those who directly experience those externalities). Except for concerns we might have about *who* receives the

⁵They also require that preferences not be distorted by things like behavioural-economic biases, and that they not be socially unacceptable, or based on habituation to unacceptable circumstances. Also they require that preferences be based on correct information, both about all of the potentially long-term costs and benefits of an impact, and all of the alternative ends that they might, upon learning of them or experiencing them, prefer to pursue. Thus, their full definition of welfare requires that preferences be not only self-interested, but also 'survive' what they call 'idealization'. I am in complete agreement with this need for idealization.

⁶For the time being, I am setting aside the universally accepted belief that a dollar is worth more to a poor person than to a wealthy person. Clearly, unless WTP estimates are adjusted by some income-dependent weights, such as those that are derived from the elasticity of marginal utility of wealth, we cannot take it to be an unbiased measure of welfare. I address this issue in section 3.

smartphone (so called distributional concerns), the reason the smartphone matters to society is that the individual gets to use it, and to the extent that their WTP for the smartphone conveys useful information about that value, it is all we need to know to evaluate the value to society of allocating the resources used to create the smartphone to the individual who consumes the phone.

We consider it to be no concern to decision makers, or the rest of society, what *kind* of value the individual gets from using the phone.⁷ There is no additional information necessary to fully capture the reason the production and consumption of the phone matters to society. It matters because the individual who consumes it says it matters, and it matters as much as they say it matters, and in colloquial terms, it is none of anyone else's business *why* they value it, and there is nothing else about the production and consumption of the phone that should matter to society.⁸ It doesn't matter whether the individual values the phone because it allows them to read the classics, or because they can play computer games. Why they value it is their business.

Another way of saying this is that the impacts about which BCA provides the appropriate kind of information are those impacts for which we want consumer sovereignty to govern, precisely because, as stated earlier, we feel it should not matter to society *why* individuals value them, and there are no other reasons they matter to society, over and above the fact that they generate value for the individuals who receive them. We can generalize from smart phones to a broader set of examples. When making decisions that affect how productive resources are allocated to goods and services, and how those goods and services are allocated to individuals, we do not want decision makers to consider the ways in which individuals use them or the kind of value they get from using them. (Unless, of course, they are using them to break the law or engage in some socially unacceptable behavior.)

Instead, we very much prefer that decision makers defer to consumers when considering the value of the goods and services they receive. Setting aside distributional concerns, we want consumers to be the sole arbiters of what is good for them. With respect to these kinds of impacts, we want consumers to be sovereign, in the very real sense that we want them to be in charge of the decision-making process. Consumer sovereignty is the cornerstone of my definition of what belongs in BCA. Because WTP preserves consumer sovereignty, only impacts for which we want consumer sovereignty to govern should be included.

Some defence is called for of the idea that the government should not be concerned with the reasons a person values a good, which is to say, the ends individuals pursue. There are two primary approaches to justifying government involvement in the decision of what ends individuals should pursue. The first is the Millian idea that there are higher and lower pleasures, which can be identified by a consensus of those who have experienced both (Mill 1999 [1859]). Thus, for example, it might matter to society whether the phone is being used to read the classics (a higher pleasure) or to play mindless games (a lower pleasure). Quantitative hedonists, such as Bentham, deny any ranking of

⁷I am ignoring socially unacceptable preferences such as child pornography.

⁸Setting aside distributional concerns, which are fundamentally different than the value of the phone itself.

the quality of pleasures, and thus insist on consumer sovereignty on the grounds that, because all ends are of equal quality, there is no way for the government to improve upon individuals' ends.⁹ However, as yet no one has demonstrated that Bentham is right and Mill wrong. But, even if we agree with Mill, there is a problem with justifying government concern with individuals' ends, as Mill himself pointed out, which is that individuals differ in their means and ability to pursue and enjoy those ends, and thus some might derive more happiness from successful pursuit of lower pleasures than from unsuccessful pursuit of higher pleasures. Mill recommended a policy of letting people figure out for themselves how to get the most out of their available means and abilities (Gibbs 1986), which is to say, consumer sovereignty.

The second approach to justifying government involvement in the ends individuals pursue is based on the idea of perfectionism. Simply put, because people care about getting value out of life, and because it is the role of government to promote that goal, the government ought to promote ways of life that are 'objectively' good (Chan 2000), even if those are not the ways of life individuals actually pursue, left to their own devices. The anti-perfectionist response is either that the only value the government should promote is autonomy, or that the government should not promote any value at all, and should leave individuals to choose their own ends, which is to say, government policy should be based on consumer sovereignty. This idea is related to the thinking of John Rawls, who argued that the only thing the government should promote is access to the resources ('primary goods') necessary for the pursuit of whatever ends individuals desire. Part of his argument is that individuals will differ in exactly how to achieve happiness or satisfaction, and have the ability to explore and revise their understanding of the ends that will allow them to do so (Kymlicka 1989). Hayek completed this argument by pointing out that the government can never have sufficient information to know what will be best for any given individual, because the level of detail necessary is lost in the aggregation inevitable in statistical reports (Hayek 1945).

There is an alternative defence of consumer sovereignty that does not involve taking a position on whether individuals know what is best for them, or whether the government knows better or should attempt to manipulate the ends they pursue. Hutt, who coined the term 'consumer sovereignty', made it clear that in his view, the reason consumers should be allowed to be sovereign is not because they know what is best for them, but because allowing them to be sovereign is 'in harmony with liberty' (Hutt 1990 [1936]: 267). Giving consumers sovereignty over how to dispose of their resources explicitly means giving them freedom of choice and autonomy. If we think of liberty, freedom, and autonomy as moral values that government ought to defend for their own sake, consumer sovereignty is the correct stance from which to defend them.

Now let's consider the kinds of impacts for which we do *not* want consumer sovereignty to govern, those for which it matters why individuals value them, or, for which considerations matter other than the effect of an impact on the

⁹Bentham (1825) famously said, 'Prejudice apart, the game of push-pin is of equal value with the arts and sciences of music and poetry'.

individual who experiences them. Consider a reduction in income inequality, an impact for which WTP has been elicited, and which should therefore be included in BCA according to the PPC. Some individuals might value a reduction in income inequality because they believe they would feel less envy and inadequacy, others because they have altruistic preferences toward the poor, and others because they believe that income inequality diminishes social solidarity. Individuals may care for multiple reasons, and will differ in which reasons matter to them, and how much they weight any given reason. Individuals' WTP will hide all of this information. Decision makers presented with aggregate WTP for a reduction in inequality will not know what kinds of value it conveys, nor how much of any given kind of value. The question at the heart of my position is, do we want decision makers to consider and weigh the different reasons that income inequality matters? If we do, then WTP does not provide the necessary information, and is thus not an appropriate measure, so income inequality, a case of distributional concerns, should not be included in BCA.

My answer is, we do want decision makers to consider and weigh the different reasons that income inequality matters. It matters to society whether individuals value a reduction in income inequality because of envy, because of altruism, or because of a concern about diminished social solidarity, and we want decision makers to weigh these values separately, in light of the value systems they believe are appropriate for governing society. For example, we might not want to make public policy on the basis of envy. By and large, we feel that individuals are best served by not making personal decision on the basis of envy. Of course, we all do make personal decisions on the basis of envy, but we might, nonetheless, not want public policy to be made on this basis. If people value the reduction because of altruism, we might consider this an appropriate goal for public policy, or we might feel that altruistic preferences ought to be satisfied through individuals' own efforts and expenditures, and not through government policy. Alternatively, we might want public decisions to benefit the poor, whether or not individuals are willing to pay for benefits to the poor on the basis of altruism. Meanwhile, if the reason for WTP for a reduction in income inequality is concern for social solidarity, we might feel this should be the domain of civil society, or alternatively, we might feel that a democratic government should act to strengthen social solidarity, because we believe democratic governance depends upon such solidarity. These are precisely the kinds of considerations we want decision makers to weigh in light of what they consider to be important.

Another way of putting this is that we don't want decisions about this kind of impact to be made on the basis of consumer sovereignty, with leaders staying out of the business of considering the kind of value involved. Rather, we want our leaders to consider the various kinds of value involved, and make decisions based on what they believe is best for society, rather than what individuals would be willing to pay for in their own private transactions. For this kind of impact, we want decision makers to apply wisdom to the consideration of why they matter to society.¹⁰

¹⁰This distinction based on consumer sovereignty applies not only to the potential Pareto criterion, but to the actual Pareto criterion as well. Consider a policy that is a Pareto improvement when taking into

To reiterate and generalize, impacts for which we want consumer sovereignty to govern are those that satisfy two criteria: i) it should not matter to society *why* individuals value them, and ii) there are no other reasons they should matter to society, over and above the fact that they matter to the individuals who receive them. These will, in general, be impacts that directly affect the individual, but do not have implications for rights (including their own, as I will explain), equity, autonomy, dignity, or other moral considerations.

One could argue that even if we don't know the reasons for people's WTP for the kinds of impacts for which we don't want consumer sovereignty to govern, their WTP should still be included. A decision maker, in addition to weighing the various reasons such impacts matter, and how much, should also have information about the total value individuals place on those impacts.¹¹ The problem with this argument is that if the values conveyed in individuals' WTP are not known, then a decision maker, weighing the pros and cons of the policy, does not know what values have already been incorporated in BCA, and thus cannot know if their separate weighing of pros and cons would involve some kind of either undercounting or double counting.

To see this, suppose a decision maker is considering a policy that involves reducing income inequality, but that generates various economic costs and benefits as well, and suppose they have a BCA that includes WTP for the reduction in inequality, using contingent valuation. Now, suppose the decision maker thinks that income inequality matters both because of the negative impact it has on the poor in particular, and because it undermines the social cohesiveness necessary for effective democratic governance. The decision maker is presumably weighing these values side by side with all of the impacts included in the BCA. Now, suppose that the main reason individuals have for being

consideration only those impacts for which we feel that consumer sovereignty should govern, but that leaves some individuals feeling that the rights of others have not been respected, to a degree that they feel offsets the gains they receive from the policy. Now consider a decision maker who cares about the impacts that warrant consumer sovereignty, and about rights. If the only information he or she is presented with is that some individuals feel they are made worse off, and not the reasons why, he or she will not know the extent to which rights have weighed in the balance for some individuals, and will thus not know how to apply his or her own judgement to the issue of rights. If the decision maker feels that the rights in question have not been violated to the same extent as the individuals who have indicated they are made worse off, or are of less moral concern that those individuals feel, he or she will not know enough to determine whether, in their judgement, the policy should be implemented. Unless the decision maker is content to base his or her decisions strictly on the referendum represented by the Pareto criterion, and does not feel responsible for applying his or her own wisdom or judgement, the Pareto criterion will not have provided the right kind of information for him or her to make a decision.

¹¹Posner and Sunstein (2017) suggest something similar, but impose the restriction that when we consider such impacts – in the case they consider, moral commitments – we should include only individuals' WTP for avoidance of mental harm caused by not having their moral commitments vindicated, and not their WTP for vindication of their moral commitments if they happen to not suffer mental harm when their moral commitments are not vindicated. (This distinction was clarified in a private communication with Eric Posner in 2019.) Although their distinction aligns with the self-interested/disinterested distinction of Adler and Posner (2006), it does not align with mine. Even though the 'mental harm' component of moral commitments might warrant consumer sovereignty, it should not, in my view, be included, because it still masks the reasons for the mental harm, which a decision maker needs to know, for the reasons I am about to give.

willing to pay for a reduction in income inequality is because of altruistic sentiments towards the poor. Presumably, then, in order to avoid over-weighting the value of altruism in their overall considerations – in other words, to avoid double counting – the decision maker should not additionally consider, to any great extent, the importance they themselves, as representatives of society as a whole, place on altruism towards the poor, and should instead add into their considerations primarily the value of social cohesion. Meanwhile, if WTP is not based on altruism, but the decision maker thinks that it is, they will under-count altruism in their weighing of different values. But there is no way to know whether it is altruism or social cohesion that has been placed upon the balance in the BCA, so the inclusion of the reduction in income inequality in the BCA has, in fact, undermined rather than improved the decision maker's ability to effectively weigh the pros and cons of the policy. It is not because how individuals feel about rights, equality, or other similar impacts doesn't matter. It is that including WTP for these things undermines the usefulness of BCA as a tool for improving the quality of public decision making.

Now, I am not arguing that the fact that individuals value impacts for which we do not want consumer sovereignty to govern should not matter to a decision maker. Of course we want decision makers to have some understanding of those things. Reducing income inequality is certainly more important in a society in which people value it a lot, or in large numbers. But WTP will not provide them with the necessary understanding, because it will only tell them *that* the impacts matter to individuals, and not *why* they matter, or how much individuals weight the various possible reasons. They will have to rely upon some other kind of information than WTP, such as surveys of public opinion. If such a survey reveals both how much people care about an issue, and why, then the decision maker has all the information they need to avoid double counting or under-counting. WTP does not provide that information, as I've explained above.

One might argue, since decision makers usually don't have access to any such alternative information about the reasons people value the kinds of impacts I argue should be excluded from benefit cost analysis, that we should at least give them *some* kind of information about the value individuals place on those impacts. By the same logic, if a baker knows that his or her customers will not get enough fish oil in their diet, perhaps he or she should add it to their cookies and cakes. My idea is that they should not, because it will destroy the value of the cookies and cakes by ruining their flavour. They should accept that the provision of fish oil is someone else's business, whether or not any such individual steps forward, and that they, the baker, are not in a position to rectify any lack of fish oil. Just so, even if we know that no one is going to conduct a well-designed survey of public opinion, it is still not our business to try to provide some proxy for the information the survey would produce, because doing so would undermine decision makers' ability to use our analysis to weigh the pros and cons of the policy. Our baker might, for good measure, place a label on his or her wares informing customers of the importance of fish oil, just as we might enumerate excluded impacts to ensure that decision makers are made aware of them. But, in my view, that is as far as a conscientious baker should go.

Steven Kelman makes an argument similar to mine in his critique of cost-benefit analysis (Kelman 1981):

[T]he efforts of economists to impute willingness to pay invariably involve bundled goods exchanged in *private* transactions.¹² Those who use figures garnered from such analysis to provide guidance for *public* decisions assume no difference between how people value certain things in private individual transactions and how they would wish those same things to be valued in public collective decisions. In making such assumptions, economists insidiously slip into their analysis an important and controversial value judgment, growing naturally out of the highly individualistic microeconomic tradition – namely, the view that there should be no difference between private behavior and the behavior we display in public social life. An alternative view – one that enjoys, I would suggest, wide resonance among citizens – would be that public, social decisions provide an opportunity to give certain things a higher valuation than we choose, for one reason or another, to [give] them in our private activities.¹³

Kelman is particularly concerned with the way that WTP *leaves out* information about the sanctity of wilderness land, or of human life, or other intangible values. My concern is not that the information is left out of WTP, but that even when WTP includes information about intangible values, it hides that information, and does not allow decision makers to consider those intangible values independently, and weigh them according to what they consider to be the appropriate value system. And at the same time, including WTP for such intangible values makes it impossible for decision makers to discern how much of the estimated net benefit of a policy consists of those values that we feel ought to be governed by consumer sovereignty, and how much represents those values for which we do not want consumer sovereignty to govern.

How does my definition of BCA differ from Adler and Posner's? First, there is an important and potentially quite large set of preferences that I treat differently from Adler and Posner, which is preferences for self-regarding rights, autonomy and dignity, which is to say an individual's own rights, autonomy and dignity. Though not mentioned explicitly in their book, by Adler and Posner's (2006) definition, BCA should include self-regarding rights, autonomy or dignity. By mine, it should not. For example, if a person has a preference for having a certain right, we cannot call that a disinterested preference, as it is a preference for something that directly affects their own life, and as such, Adler and Posner would include it in BCA.¹⁴ However, I would not call it a sovereignty-warranting preference. I have said that a policy impact warrants consumer sovereignty if the reason it matters to the individual who experiences it does not matter to society,

¹²And I would add that when, instead, WTP is elicited with a contingent valuation survey, every effort is made to frame the hypothetical transaction as a private one.

¹³And, in my view, to give certain other things a lower valuation, such as the desire to ban transgender individuals from the military (for which I have elicited WTP) because of a feeling that they are inherently not worthy of the same rights as others (one of the reasons considered somewhat or very important by those respondents willing to pay for the ban).

¹⁴I have confirmed this in a private communication with Adler.

and there are no other reasons that it should matter to society. Self-regarding rights do not satisfy the second part of this criterion. Your right to a thing is valuable to society over and above the value you place on it. To see this, consider the right of minorities to vote. Suppose, for some reason, that every single member of a given minority did not want to vote, and did not care if they had the right to vote. I believe most people would agree that it would still matter to society whether or not they were granted the right to vote. The same should be true of autonomy and dignity, and potentially other impacts that individuals value in a self-regarding sense, but that matter to society for additional reasons.

In addition, there is a more general way in which my definition of BCA differs from Adler and Posner's. They adhere to the traditional position in Welfare Economics that BCA is a measure of welfare (Adler and Posner 2006). This is problematic because there is disagreement about how welfare should be defined. There are three main approaches. The first is experientialist: a person's welfare is high or low to the extent that their experience is pleasant or unpleasant, satisfying or unsatisfying, etc. The second is objectivist: a person's welfare is high or low to the extent that they have acquired or attained certain objectively determined dimensions of a 'good' life. The third is preferentialist: a person's welfare is high or low to the extent that their preferences are satisfied, whatever their preferences may be based on. (Though a preferentialist might restrict the set of preferences that should be considered part of welfare, as Adler and Posner do.)

Now, consider an individual who has an average income for which they do not have to work, and who has high willingness to pay for a smartphone because it allows them to fill all of their time with mindless computer games, despite having full information of the value of other alternatives, perfect rationality, and having no second-order preference for not playing games. By a hedonic account of welfare, the individual's high willingness to pay reflects high welfare only if the individual feels good while playing on their phone. Thus, whether their WTP should be included in BCA is ambiguous. To an objectivist, the individual's high willingness to pay does not reflect high welfare, because playing mindless games is not typically thought of as part of an objectively good life. Thus, their WTP should not be included in BCA. To a preferentialist, the person's welfare is high by definition, because their willingness to pay reflects a preference for it, and thus their WTP should be included in BCA.

Any welfarist definition of BCA requires taking a position on the definition of welfare. By contrast, my definition of BCA remains agnostic as to the definition of welfare, and thus does not require taking a position in this intractable debate, precisely because it bases the choice of whether to include an impact in BCA purely on whether it warrants consumer sovereignty, and not on whether it should count as welfare. By my definition, regardless of how one thinks welfare should be defined, the WTP of the player of mindless games should be included in BCA because it warrants consumer sovereignty.

My approach to BCA should be preferred by both objectivists and restricted preferentialists like Adler and Posner for two reasons. First, because, as I've explained above, consumer sovereignty should be valued for reasons other than the maximization of any particular definition of welfare, regardless of how that word is defined. There is evidence that both camps do in fact endorse consumer

sovereignty, either implicitly, by endorsing the values that consumer sovereignty promotes, or else explicitly. To see this, note that the values promoted by consumer sovereignty are included in more than one list of objective goods in the objectivist literature. For example, Griffin (1998) explicitly includes autonomy, and Nussbaum (2001) includes control over one's environment, of which autonomy is surely a part. If autonomy is viewed as being of fundamental importance, then consumer sovereignty must also be viewed as being of fundamental importance. Meanwhile, Adler and Posner (2006: 32), in their rejection of objectivism, explicitly endorse consumer sovereignty when they say, '[A]ll of these [objectivist] accounts overlook the crucial point that each individual is a (partial) sovereign with respect to his own welfare'. Respect for consumer sovereignty is fundamental to their defence of preferentialism. However defensible one's definition of welfare may be, adopting a definition of BCA that is at odds with a value one considers to be of fundamental importance is inconsistent.

The second reason that objectivists and restricted preferentialists should prefer my definition of BCA is that the restriction I impose on what should be included in BCA ensures that it will remain useful and informative to decision makers. Without my restriction, the usefulness of BCA is undermined, owing to the problem of over- and under-counting of non-sovereignty-warranting impacts, which arises from the fact that WTP gives us no information about why individuals value impacts, making it impossible for decision makers to know how much weight to place on any non-sovereignty-warranting values involved in a policy. If, in any particular BCA, the objectivist or the restricted preferentialist feel that the tally of WTP in a BCA is not reflective of what they think of as welfare, they can add a comment to this effect, as is common practice whenever there are policy impacts about which we feel BCA does not provide adequate information. By contrast, the inclusion of impacts that do not warrant consumer sovereignty cannot be adequately addressed with this kind of caveat, because there is no way at all of determining the reasons individuals have for valuing an impact, and thus no way of eliminating the problem of over- or under-counting.¹⁵ However strongly one believes in the importance of measuring and communicating aggregate welfare, restricting what is included in BCA on the basis of some definition of welfare, rather than on the basis of consumer sovereignty, does harm to BCA itself.

A challenge arises if a good is valuable to an individual at least in part because it allows them to exercise a right. In this case, the individual's WTP for that good would capture, at least in part, their preference for a self-regarding right. For example, an individual might have WTP for a bus ride at least in part because it allows them to get to a polling station to exercise their right to vote, or an individual might value a smartphone because it allows them to pursue an education, which some consider a right. Under my definition of BCA, we would need to decompose the individual's WTP into WTP for a self-regarding right, and WTP for whatever other, sovereignty-warranting outcome they might get from the bus ride. In practice, it might be possible to attempt this decomposition using a well-designed contingent valuation survey. If impacts are

¹⁵I expand on this idea in section 3, in which I compare the cost of including what should not be included to the cost of excluding what should.

monetized using market data, however, the decomposition would almost certainly be impracticable. Solving this practical problem is beyond the scope of this paper.

3. The importance of avoiding incorrect inclusion

In the preceding section I've presented a conceptual definition of what should and should not be included in BCA, arguing that any policy impact that does not warrant consumer sovereignty should not be included in BCA, and vice versa. In practice, the distinction may not be obvious in all cases, so it is important to consider whether we should err on the side of excluding things that warrant consumer sovereignty, or on the side of including things that do not. I argue for the former. We should be particularly careful to avoid type II error, incorrectly including impacts for which we do not want consumer sovereignty to govern (false positives), at some potential cost in terms of type I error, incorrectly excluding things for which we do (false negatives).

The reason why is this. In terms of the appropriateness of BCA as a tool for informing public decision making, the cost of including what should be excluded is greater than the cost of excluding what should be included. What is the cost of including what should be excluded? The answer lies, to some extent, in the fact that BCA not only uses WTP to measure and convey social value, but also aggregates WTP, first across the affected population for any given impact, and then across impacts. If all of the impacts are of the type for which we want consumer sovereignty to govern, this aggregation does not reduce the appropriateness of the information, because we do not care about any of the reasons that any of the impacts matter, to individuals or society as a whole. All we care about is the total amount of value individuals say they get, on net. We can aggregate value across individuals and across impacts without any loss of information, because all of the impacts are those for which we feel consumer sovereignty should govern.

To see the cost of type II error, consider what might happen if, instead of including only impacts over which we want consumer sovereignty to govern, we include some for which we do not. Then, not only does the information about the value individuals place on that impact not convey the kind of information we want decision makers to have, but, because of the aggregation inherent in BCA, we no longer know what our measure of net benefits actually does convey. It is a muddle of information about the value of individuals' own direct experience of impacts, and information about how they feel about the various other values they think are affected by one or more of the impacts. Decision makers, looking at our measure of net benefits, do not know how much of it represents values that they ought to give no further consideration to, and how much represents values that they ought to consider themselves as being tasked with weighing, according to their own sense of what is wise. In a sense, this would be akin to adding gallons of oil and gallons of milk to come up with an aggregate measure called, simply, 'gallons'. Except for a very narrow range of purposes, this would be an uninformative measure. If a decision maker chose to

treat it as if it were informative, it could distort their decisions, leading them to make decisions that they would not have made if there had been no type II error.

As I discussed earlier, the basic idea is that if a particular kind of value has already been included in BCA, the decision maker, not knowing this, might err, either by additionally weighing that value to the extent they consider appropriate, which would be a kind of double counting, or, if they know that the value in question has been included in BCA, by concluding that they do not need to additionally weigh it according to their own sense of how much it matters, which could lead to undercounting. In either case, their decision might be different than if BCA had contained only the kind of value I am proposing, leaving them to weigh other values exactly as they consider appropriate. This concern applies equally to my definition of BCA and to Adler and Posner's. With the exception of self-regarding rights, the self-interested preferences they include ought to matter to society solely because they matter to the individual, whereas disinterested preferences, which can be thought of as those that touch upon values that matter to society as a whole, should be weighed independently, in order to avoid over- or under-counting.

One might argue that disaggregating WTP for different kinds of impacts would salvage the value of WTP as a way to convey *some* information about the value of impacts for which we do not want consumer sovereignty to govern. But this goes against the basic purpose of BCA, which is to compute a single measure of net benefit. And anyway, as I explained earlier, even a disaggregated measure of WTP for the kinds of impacts or values I believe should be included would not provide the decision maker with useful information, because they would still not know what specific values were actually contained in the measure of WTP.

What about the cost of type I error, excluding from BCA impacts for which we arguably might want consumer sovereignty to govern? We have a well-established way of compensating for this kind of omission. We know how to enumerate impacts that should be included but are not part of our quantitative measure of net benefits. That is precisely what *Circular A-4* calls for. In the case of impacts excluded from BCA incorrectly, we can simply enumerate them and weigh them as part of our overall tally of net benefit. We already do this. The cost of type I error is low. So we should err on the side of excluding things that ought to be included in order to ensure that we correctly exclude everything that should not. In future work I will develop an operationalizable rule for implementing my definition that satisfies this desideratum.

A final note must be made about the seemingly unavoidable bias in WTP as a measure of value, and the way this threatens to violate the insistence on avoiding type II error. Because of the widely held belief that a dollar is worth more to a poor person than to a wealthy person – due to so-called ‘diminishing marginal utility of wealth’ – it is typically accepted that if a wealthy person and a poor person both place a value of \$1,000 on a smartphone, it is in fact of greater value to the poor person, because each dollar matters more to them, and thus expresses more value. As a result, using WTP as a measure of value unavoidably results in the inclusion of information we do not want included in BCA: in addition to the value generated by policy impacts, WTP contains information about the wealth of the individuals who receive those impacts. But of course we can never know what part of WTP represents value, and what part represents income. As such,

by using WTP as a measure of value, we are automatically incurring the cost of type II error. And, as I have argued, this cost is high.

Thus, to implement either my definition, or Adler and Posner's definition of BCA correctly, it would be necessary to 'adjust' WTP estimates on the basis of wealth so that it would not be a biased measure of value, and so that BCA would not contain information it ought not to contain. There has been considerable work on the development of so-called 'efficiency weights', well summarized by Cowell and Gardiner (2000). Though any practicable approach to determining weights will be fraught with theoretical and measurement challenges, an approach based on estimates of the elasticity of marginal utility of wealth from risk preferences has been developed, and is in use in the UK (HM Treasury 2018). In a very real sense, in order for BCA to satisfy my definition, WTP estimates must be adjusted using this approach or some other. For those who consider this kind of adjustment unsatisfactory, this problem of incorrect inclusion will not be fully solvable, and the claim that BCA is a measure of sovereignty-warranting preferences will not be fully valid.

4. Conclusion

To summarize, there is no way to rectify the idea that BCA is defined by the potential Pareto criterion, and the idea that it should not include some of the kinds of impacts we typically think it should not include. We must abandon one or other of these ideas. If we abandon the idea that certain kinds of impacts should not be included, effectively doubling down on the PPC, then in my view we render BCA un-useful for communicating about the pros and cons of policies, in the sense that the information it contains is no longer the appropriate kind of information. Thus, like Adler and Posner, I believe we should abandon the PPC definition of BCA. If we do, then we need a new definition. My definition is that the only impacts that should be included in BCA are those for which we want consumer sovereignty to govern, because WTP, as a measure of value, preserves consumer sovereignty, and thus only provides useful information about such impacts. I have also argued that we should be particularly concerned about excluding impacts for which we do not want consumer sovereignty to govern, perhaps at some cost in terms of excluding some for which we do.

Abandoning the potential Pareto criterion definition of BCA would be a non-trivial step. Not only does the PPC definition have tremendous resonance, and a long and proud pedigree, it is also, I would argue, easier to teach than either my definition or Adler and Posner's. The same case could have been made for clinging to Newton's laws in favour of the theory of relativity, but this path was ultimately not chosen, for reasons similar to mine. And yet, at the same time, just as Newton's laws remain a relevant tool for understanding the physical universe, both as an introductory lesson, and as a reasonable approximation when things are not moving at near the speed of light, I believe that the PPC will remain a useful concept for introducing the basic idea of net benefit, and conceptualizing the basic enterprise of BCA. I simply think that when it comes to correctly teaching BCA, and even more when it comes to correctly practicing

BCA, we need to move beyond the PPC to something better. I believe I have provided a way to do this.

Acknowledgements. I am grateful to Matthew Adler, Eric Posner, Eugene Smolensky, Daniel Baker, Lee Friedman, Rob Letzler, Rob Moore, and attendees of the 2019 conference of the Society for Benefit Cost Analysis. I would also like to acknowledge the extremely helpful comments of anonymous reviewers.

References

- Adler M.D. and E.A. Posner** 2006. *New Foundations of Cost-benefit Analysis*. Cambridge, MA: Harvard University Press.
- Bentham J.** 1825. *The Rationale of Reward*. John and HL Hunt.
- Boardman A.E., D.H. Greenberg, A.R. Vining and D.L. Weimer** 2011. *Cost-benefit Analysis: Concepts and Practice*. 4th edition. Upper Saddle River, NJ: Prentice Hall.
- Bostani M. and A. Malekpoor** 2012. Critical analysis of Kaldor-Hicks efficiency criterion, with respect to moral values, social policy making and incoherence. *Advances in Environmental Biology* **6**, 2032–2038.
- Chan J.** 2000. Legitimacy, unanimity, and perfectionism. *Philosophy & Public Affairs* **29**, 5–42.
- Cowell F. and K. Gardiner** 2000. *Welfare Weights*. London: Office of Fair Trading.
- Gibbs B.** 1986. Higher and lower pleasures. *Philosophy* **61**, 31–59.
- Gramlich E.** 1990. *A Guide to Benefit-Cost Analysis*. Long Grove, IL: Waveland Press.
- Griffin J.** 1998. *Value Judgement: Improving our Ethical Beliefs*. Oxford: Oxford University Press.
- Hayek F.A.** 1945. The use of knowledge in society. *American Economic Review* **35**, 519–530.
- HM Treasury** 2018. *The Green Book: Central Government Guidance on Appraisal and Evaluation*.
- Hicks J.R.** 1941. The rehabilitation of consumers' surplus. *Review of Economic Studies* **8**, 108–116.
- Hotelling H.** 1938. The general welfare in relation to problems of taxation and of railway and utility rates. *Econometrica* **6**, 242–269.
- Hutt W.H.** 1990 [1936]. *Economists and the Public: A Study of Competition and Opinion*. Piscataway, NJ: Transaction Publishers.
- Kaldor N.** 1939. Welfare propositions of economics and interpersonal comparisons of utility. *Economic Journal* **49**, 549–552.
- Kelman S.** 1981. Cost-benefit analysis: an ethical critique. *Regulation* **5**, 33.
- Kymlicka W.** 1989. Liberal individualism and liberal neutrality. *Ethics* **99**, 883–905.
- Mill J.S.** 1999 [1859]. *On Liberty*, ed. E. Alexander. Peterborough, ON: Broadview Press.
- Nussbaum M.C.** 2001. *Women and Human Development: The Capabilities Approach (Vol. 3)*. Cambridge: Cambridge University Press.
- Office of Management and Budget.** 2003. *Circular A-4, Regulatory Analysis*.
- Posner E.A. and C.R. Sunstein** 2017. Moral commitments in cost-benefit analysis. *Virginia Law Review* **103**, 1809–1860.
- Zerbe R.O.** 2004. Should moral sentiments be incorporated into benefit-cost analysis? An example of long-term discounting. *Policy Sciences* **37**, 305–318.

Daniel Acland is on the faculty of the Goldman School of Public Policy at the University of California, Berkeley. He has taught How to Conduct Benefit Cost Analysis since 2010, and has more recently developed an alternative course named 'Principles of Benefit Cost Analysis for Policy Makers', from which his ideas on the Economic Philosophy of Benefit Cost Analysis and Welfare Measurement have emerged. He also conducts research on the Economic Philosophy of Behavioural Economics, and at the intersection of Benefit Cost Analysis and Behaviour Economics.

URL: <https://gspp.berkeley.edu/directories/faculty/dan-acland>

Cite this article: Acland D (2022). What's in, what's out? Towards a rigorous definition of the boundaries of benefit-cost analysis. *Economics & Philosophy* **38**, 34–50. <https://doi.org/10.1017/S0266267120000486>